

Figure 1A

1	CACCAGCAGTAGTAGCAGAAGCGAAGAGCGCAAACGCAACCGCTCTCCCCGCGGTTGGC	60
61	CGATTCATTAATGCAGCTGGCAGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCA	120
121	ACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTC	180
181	CGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATG	240
241	ACCATGATTACGCCAAGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCC	300
301	ACCGCGGTGGCGGCCGCTCTAGAACTAGTGATCCCCCGGGCTGCAGGAATTCGGCACGA	360
361	GAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCAT	420
1		M 1
421	GGCTTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGG	480
2	A S L G Q I L F W S I I S I I I I L A G	21
481	AGCAATTGCACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTAC	540
22	A I A L I I G F G I S G R H S I T V T T	41
541	TGTCGCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGA	600
42	V A S A G N I G E D G I L S C T F E P D	61
601	CATCAAACCTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCA	660
62	I K L S D I V I Q W L K E G V L G L V H	81
661	TGAGTTCAAAGAAGGCAAAGATGAGCTGTCCGAGCAGGATGAAATGTTTCAGAGGCCGGAC	720
82	E F K E G K D E L S E Q D E M F R G R T	101
721	AGCAGTGTTTGCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCA	780
102	A V F A D Q V I V G N A S L R L K N V Q	121
781	ACTCACAGATGCTGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGC	840
122	L T D A G T Y K C Y I I T S K G K G N A	141
841	TAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGC	900
142	N L E Y K T G A F S M P E V N V D Y N A	161
901	CAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTG	960
162	S S E T L R C E A P R W F P Q P T V V W	181
961	GGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCT	1020
182	A S Q V D Q G A N F S E V S N T S F E L	201

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1B

1021	GAAC TCTGAGAATGTGACCATGAAGTTGTGTCTGTGCTCTACAATGTTACGATCAACAA	1080
202	N S E N V T M K V V S V L Y N V T I N N	221
1081	CACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTGAC	1140
222	T Y S C M I E N D I A K A T G D I K V T	241
1141	AGAATCGGAGATCAAAAGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTG	1200
242	E S E I K R R S H L Q L L N S K A S L C	261
1201	TGTCTCTTCTTTCTTGGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCT	1260
262	V S S F F A I S W A L L P L S P Y L M L	281
1261	AAAATAATGTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTAC	1320
282	K *	283
1321	AGAACTATTTACCACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCAT	1380
1381	ATCTAGAAGTCTGGAGTGAGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAG	1440
1441	GCTCCAATATGAACAAGATAAACTCTATCTTCAAAGACATATTAGAAGTTGGGAAAATAAT	1500
1501	TCATGTGAAGTAGACAAGTGTGTTAAGAGTGATAAGTAAATGCACGTGGAGACAAGTGC	1560
1561	ATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGGAGTGAGAGGACAGGATAG	1620
1621	TGCATGTTCTTTGTCTCTGAATTTTGTAGTTATATGTGCTGTAATGTTGCTCTGAGGAAGC	1680
1681	CCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCACAAATTAAGCTGTAGTAT	1740
1741	GTACCCTAAGACGCTGCTAATCGACTGCCACTTCGCAACTCAGGGGCGGCTGCATTTTAG	1800
1801	TAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTCTCTTCCC	1860
1861	AACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGTCG	1920
1921	GCGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAACAAATGCGGGTTT	1980
1981	ATTTCTCAGATGATGTTTCATCCGTGAATGGTCCAGGGAAGGACCTTTCACCTTGACTATA	2040
2041	TGGCATTATGTCATCACAAAGCTCTGAGGCTTCTCCTTTCCATCCTGCGTGACAGCTAAG	2100
2101	ACCTCAGTTTTCAATAGCATCTAGAGCAGTGGGACTCAGCTGGGGTGATTTGCCCCCCA	2160
2161	TCTCCGGGGGAATGTCTGAAGACAATTTTGGTTACCTCAATGAGGGAGTGGAGGAGGATA	2220

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted April 28, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1C

2221 CAGTGCTACTACCAACTAGTGGATAAAGGCCAGGGATGCTGCTCAACCTCCTACCATGTA 2280  
2281 CAGGACGTCTCCCCATTACAACCTACCCAATCCGAAGTGTCAACTGTGTCAGGACTAAGAA 2340  
2341 ACCCTGGTTTTGAGTAGAAAAGGGCCTGGAAAGAGGGGAGCCAACAAATCTGTCTGCTTC 2400  
2401 CTCACATTAGTCATTGGCAAATAAGCATTCTGTCTCTTTGGCTGCTGCCTCAGCACAGAG 2460  
2461 AGCCAGAACTCTATCGGGCACCAGGATAACATCTCTCAGTGAACAGAGTTGACAAGGCCT 2520  
2521 ATGGGAAATGCCTGATGGGATTATCTTCAGCTTGTGAGCTTCTAAGTTCTTTCCCTTC 2580  
2581 ATTCTACCCTGCAAGCCAAGTTCTGTAAGAGAAATGCCTGAGTTCTAGCTCAGGTTTTCT 2640  
2641 TACTCTGAATTTAGATCTCCAGACCCCTTCCTGGCCACAATTCAAATTAAGGCAACAAACA 2700  
2701 TATACCTTCCATGAAGCACACACAGACTTTTGAAAGCAAGGACAATGACTGCTTGAATTG 2760  
2761 AGGCCTTGAGGAATGAAGCTTTGAAGGAAAAGAATACTTTGTTTCCAGCCCCCTTCCCAC 2820  
2821 ACTCTTCATGTGTTAACCACATGCCTTCCTGGACCTTGAGCCACGGTGACTGTATTACAT 2880  
2881 GTTGTTATAGAAAACCTGATTTTAGAGTTCTGATCGTTCAAGAGAATGATTAAATATACAT 2940  
2941 TTCTTAAAAAAAAAAAAAAAAAACTCGAGGGGGGGCCCGGTACCCAATTGCCCCATAGT 3000  
3001 GAGTCGTATTACAATTCACCTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGC 3060  
3061 GTTACCCAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAA 3120  
3121 GAGGCCCCGACCGATCGCCCTTCCCAACAATTGCGCAGCCTGAATGGCGAATGGCAAATT 3180  
3181 GTAAGCGTTAATATTTTGTAAAAATTGCGGTTAAATTTTGTAAATCAGCTCATTTTTT 3240  
3241 AACCAATAGGCCGAAATCGGC AAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGG 3300  
3301 TTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGTGTTCACCGCGGTGA 3357

Figure 2

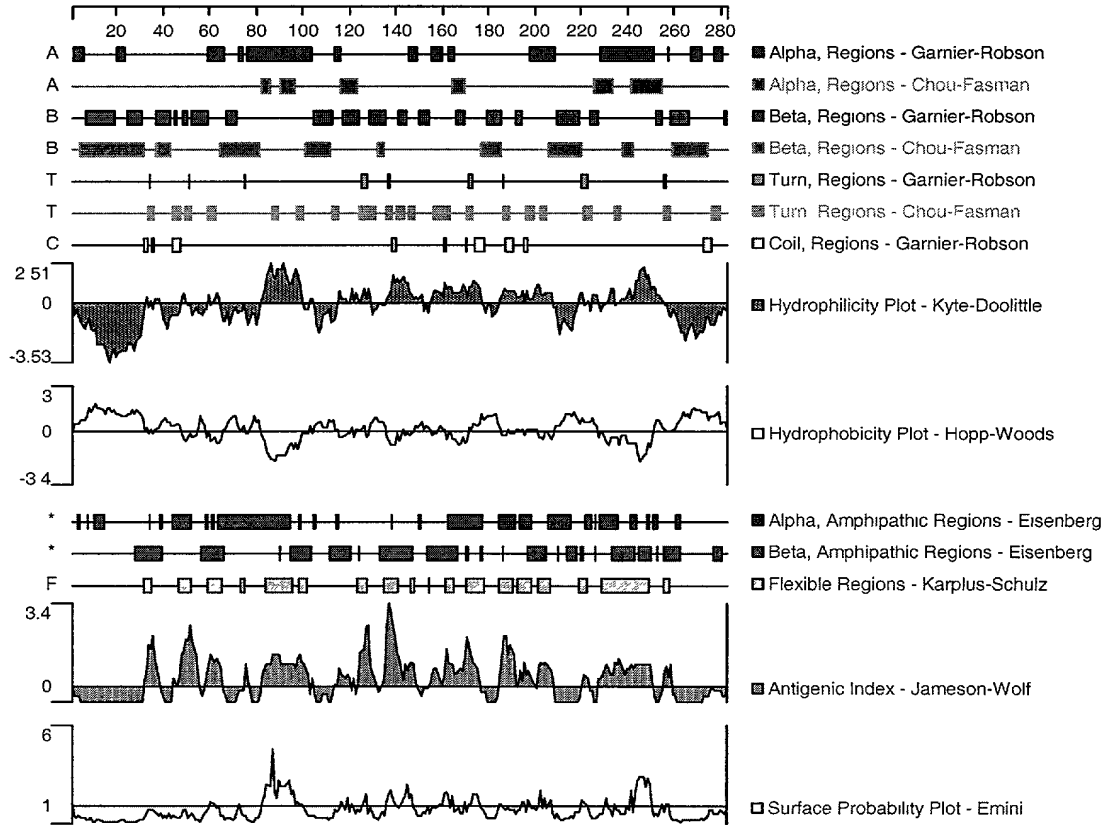


Figure 3A

1	CCACGCGTCCGGAATGAACAAC	TTTCTCTCTTGAATATATCTTAACGCCAAATTTTGA	60
61	GTGCTTTTTTGT	TACCCATCCTCATATGTCCCAGCTGGAAAGAATCCTGGGTGGAGCTA	120
121	CTGCATGTTGATTGTTTTGTTTTTCCTTTTGGCTGTT	CATTTTGGTGGCTACTATAAGGA	180
181	AATCTAACACAAACAGCAACTGTTTTTGTGTTTACTTTTGCATCTTTACTTGTGGAGC		240
241	TGTGGCAAGTCCTCATATCAAATACAGAACATGATCTTCCTCCTGCTAATGTTGAGCCTG		300
1		M I F L L L M L S L	10
301	GAATTCAGCTTCACCAGATAGCAGCTTTATTCACAGTGACAGTCCCTAAGGAAC	TGTAC	360
11	E L Q L H Q I A A L F T V T V P K E L Y		30
361	ATAATAGAGCATGGCAGCAATGTGACCCTGGAATGCAACTTTGACACTGGAAGTCATGTG		420
31	I I E H G S N V T L E C N F D T G S H V		50
421	AACCTTGAGCAATAACAGCCAGTTTGCAAAAGGTGGAAATGATACATCCCCACACCGT		480
51	N L G A I T A S L Q K V E N D T S P H R		70
481	GAAAGAGCCACTTTGCTGGAGGAGCAGCTGCCCTAGGGAGGCCTCGTTCCACATACCT		540
71	E R A T L L E E Q L P L G K A S F H I P		90
541	CAAGTCCAAGTGAGGGACGAAGGACAGTACCAATGCATAATCATCTATGGGGTCGCCTGG		600
91	Q V Q V R D E G Q Y Q C I I I Y G V A W		110
601	GACTACAAGTACCTGACTCTGAAAGTCAAAGCTTCTACAGGAAAATAAACTCACATC		660
111	D Y K Y L T L K V K A S Y R K I N T H I		130
661	CTAAAGGTTCCAGAAACAGATGAGGTAGAGCTCACCTGCCAGGCTACAGGTTATCCTCTG		720
131	L K V P E T D E V E L T C Q A T G Y P L		150
721	GCAGAAGTATCCTGGCCAAACGTCAGCGTTCTTGCCAACACCAGCCACTCCAGGACCCCT		780
151	A E V S W P N V S V P A N T S H S R T P		170
781	GAAGGCCTCTACCAGGTCACCAAGTGTCTGCGCCTAAAGCCACCCCTGGCAGAACTTC		840
171	E G L Y Q V T S V L R L K P P P G R N F		190
841	AGCTGTGTGTTCTGGAATACTCACGTGAGGGAAC	TACTTTGGCCAGCATTGACCTTCAA	900
191	S C V F W N T H V R E L T L A S I D L Q		210
901	AGTCAGATGGAACCCAGGACCCATCCAAC	TGGCTGCTTCACATTTTCATCCCCTCCTGC	960
211	S Q M E P R T H P T W L L H I F I P S C		230
961	ATCATTGCTTTTCATTTTCATAGCCACAGTGATAGCCCTAAGAAAACAACTCTGTCAAAG		1020
231	I I A F I F I A T V I A L R K Q L C Q K		250

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3B

1021	CTGTATTCTTCAAAGACACAACAAAAGACCTGTCACCACAACAAAGAGGGAAGTGAAC	1080
251	L Y S S K D T T K R P V T T T K R E V N	270
1081	AGTGCTGTGAATCTGAACCTGTGGTCTTGGGAGCCAGGGTGACCTGATATGACATCTAAA	1140
271	S A V N L N L W S W E P G *	284
1141	GAAGCTTCTGGACTCTGAACAAGAATTCGGTGGCCTGCAGAGCTTGCCATTGCACTTTT	1200
1201	CAAATGCCTTTGGATGACCCAGCACTTTAATCTGAAACCTGCAACAAGACTAGCCAACAC	1260
1261	CTGGCCATGAAACTTGCCCCTTCACTGATCTGGACTCACCTCTGGAGCCTATGGCTTTAA	1320
1321	GCAAGCACTACTGCACTTTACAGAATTACCCCACTGGATCCTGGACCCACAGAATTCCTT	1380
1381	CAGGATCCTTCTTGCTGCCAGACTGAAAGCAAAGGAATTATTTCCCTCAAGTTTTCTA	1440
1441	AGTGATTTCAAAAGCAGAGGTGTGTGGAATTTCCAGTAACAGAAACAGATGGGTGCC	1500
1501	AATAGAGTTATTTTTATCTATAGCTTCCTCTGGGTACTAGAAGAGGCTATTGAGACTAT	1560
1561	GAGCTCACAGACAGGGCTTCGCACAAACTCAAATCATAATTGACATGTTTTATGGATTAC	1620
1621	TGGAATCTTGATAGCATAATGAAGTTGTTCTAATTAACAGAGAGCATTAAATATACACT	1680
1681	AAGTGACAAATTGTGGAGTAAAGTCATCAAGCTCTGTTTTTGAGGTCTAAGTCACAAAG	1740
1741	CATTTGTTTTAACCTGTAATGGCACCATGTTAATGGTGGTTTTTTTTTGAACACATC	1800
1801	TTTCCTTTAAAAATTATTGGTTTCTTTTTATTGTTTTTACCTTAGAAATCAATTATATA	1860
1861	CAGTCAAAAATATTTGATATGCTCATACGTTGTATCTGCAGCAATTCAGATAAGTAGCT	1920
1921	AAAATGGCCAAAGCCCCAAACTAAGCCTCCTTTTCTGGCCCTCAATATGACTTTAAATTT	1980
1981	GACTTTTCAGTGCCTCAGTTTGACATCTGTAATACAGCAATGCTAAGTAGTCAAGGCCT	2040
2041	TTGATAATTGGCACTATGGAAATCCTGCAAGATCCCACTACATATGTGTGGAGCAGAAGG	2100
2101	GTAACTCGGCTACAGTAACAGCTTAATTTTGTTAAATTTGTTCTTTATACTGGAGCCATG	2160
2161	AAGCTCAGAGCATTAGCTGACCCTTGAACATATCAAATGGGCACATTAGCTAGTATAACA	2220
2221	GACTTACATAGGTGGGCCTAAAGCAAGCTCCTTAACTGAGCAAAATTTGGGGCTTATGAG	2280

**Figure 3C**

```

2281  AATGAAAGGGTGTGAAATTGACTAACAGACAAATCATACATCTCAGTTTCTCAATTCTCA      2340
2341  TGTAAATCAGAGAATGCCTTTAAAGAATAAAACTCAATTGTTATTCTTCAAAAAAAAAAA      2400
2401  AAAAAA      2406

```

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)	(49)	(50)	(51)	(52)	(53)	(54)	(55)	(56)	(57)	(58)	(59)	(60)	(61)	(62)	(63)	(64)	(65)	(66)	(67)	(68)	(69)	(70)	(71)	(72)	(73)	(74)	(75)	(76)	(77)	(78)	(79)	(80)	(81)	(82)	(83)	(84)	(85)	(86)	(87)	(88)	(89)	(90)	(91)	(92)	(93)	(94)	(95)	(96)	(97)	(98)	(99)	(100)
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	

Figure 4

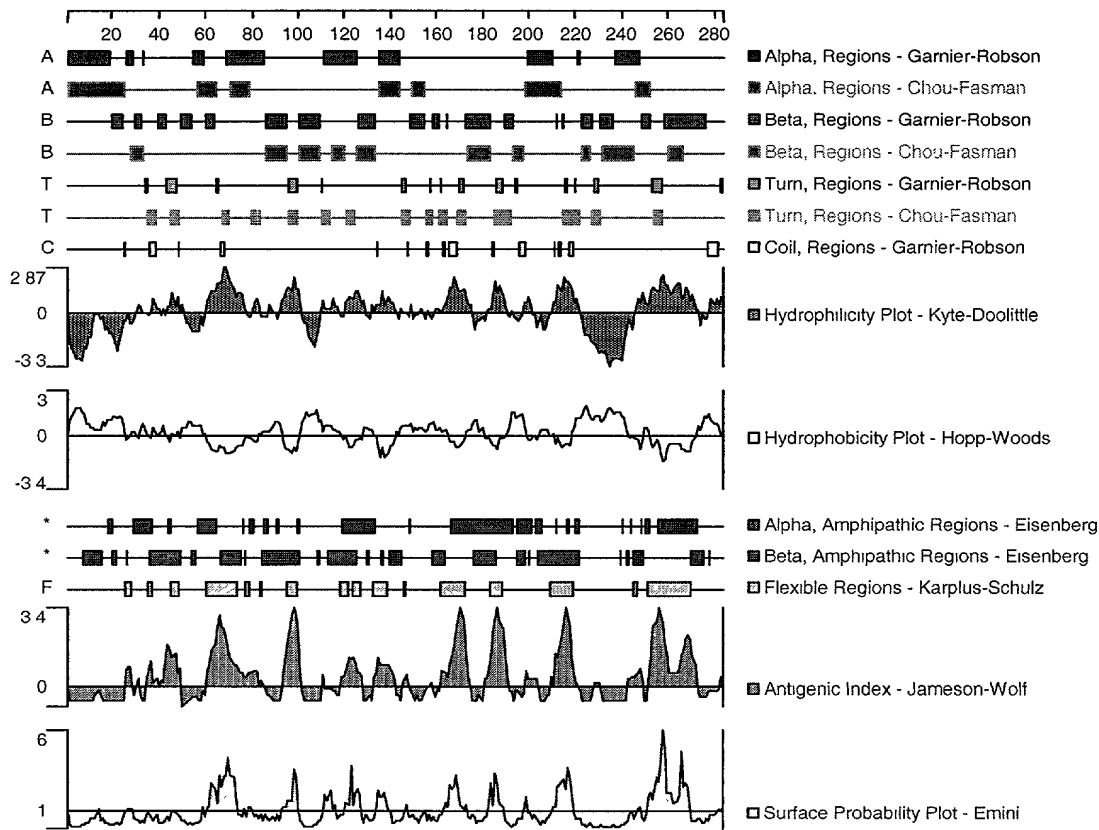




Figure 5A

1	GGCACGAGCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCT	60
1	M A	2
61	CTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTT	120
3	L M L S L V L S L L K L G S G Q W Q V F	22
121	GGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGAGGACGCAGCATTCTCCTGTTTCCTG	180
23	G P D K P V Q A L V G E D A A F S C F L	42
181	TCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCAGTTCCTCAGGGGCCAGTTCCTCTAGC	240
43	S P K T N A E A M E V R F F R G Q F S S	62
241	GTGGTCCACCTCTACAGGGACGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAA	300
63	V V H L Y R D G K D Q P F M Q M P Q Y Q	82
301	GGCAGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTCTGAGGCTGGAA	360
83	G R T K L V K D S I A E G R I S L R L E	102
361	AACATTACTGTGTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCTCAGTCTTACTAC	420
103	N I T V L D A G L Y G C R I S S Q S Y Y	122
421	CAGAAGGCCATCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCCTCTCATTTCCATC	480
123	Q K A I W E L Q V S A L G S V P L I S I	142
481	GCGGGATATGTTGATAGAGACATCCAGCTACTCTGTCTCAGTCCTCGGGCTGGTTCCCCCGG	540
143	A G Y V D R D I Q L L C Q S S G W F P R	162
541	CCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTTGTCCACAGACTCCAGGACAAAC	600
163	P T A K W K G P Q G Q D L S T D S R T N	182
601	AGAGACATGCATGGCCTGTTTGTATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGGG	660
183	R D M H G L F D V E I S L T V Q E N A G	202
661	AGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAG	720
203	S I S C S M R H A H L S R E V E S R V Q	222
721	ATAGGAGACTGGAGAAGAAAGCACGGACAGGCAGGTAAAAGAAAATATTCCTCTTCACAC	780
223	I G D W R R K H G Q A G K R K Y S S S H	242
781	ATTTATGACTCCTTTCCAAGTCTCTCGTTTATGGATTTTATATCCTGAGGCCCGTGGGT	840
243	I Y D S F P S L S F M D F Y I L R P V G	262
841	CCCTGCAGAGCCAAGCTTGTGATGGGAACCTCTGAAATTGCAGATTCTGGGGGAGGTGCAT	900
263	P C R A K L V M G T L K L Q I L G E V H	282
901	TTTGTAGAGAAGCCCCATAGCCTTCTTCAGATCTCTGGAGGGTCCACAACACTCAAAAAG	960
283	F V E K P H S L L Q I S G G S T T L K K	302

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5B

961	GGTCCCAATCCTTGGTCTTTCCCTTCTCCCTGCGCCCTGTTTCCCACGTGAGCACGGAAC	1020
303	G P N P W S F P S P C A L F P T *	319
1021	TGCCTGCTCTCTCTGCTTGTTCAGAATTGAGAGACGCCCCGAAACACGCAGGTACCAA	1080
1081	CGCCTGAGAGGGTAACAGTGGGCATGGAGTAGGAAGATGACCAGTGACAGATATGGAGCC	1140
1141	CATCCAGCTTGTAGACAGCAAATCTGTGATGCCCGAATCCACCCCAGGGTGCAGCTGCCT	1200
1201	CTAAATACACTTCTTGGCCCAGGACTTGGAGGGAAAAGCGTAGGGACTGGGTCAGCTAGG	1260
1261	AGGGGTCACAGGCAAGACGCCAGGGAAGTGGAGGCATTAGTAGCTGGCTTCTAGGGGTCT	1320
1321	GTGCAAAGGGGAACGAAGTGAAGTTAGCAGGAACTGGTGGGTGGAAGGAAGCTGAATCCT	1380
1381	GGAGTCACTCAAGGTCTCACAAAGTCAAATAGAGGGCTTACGTGGGAGGGCAGTGGTAGG	1440
1441	GCTGGGTGAACATCTCATGGTTGAGCATCTCCAAGCATCAGTGAGGCACGGGGCTGCCC	1500
1501	TGGAGAAGGTACATGGCTGGTGGGATAGTGGGACTGGCCGGATCCTACCCGGAGCCAGTC	1560
1561	TGCAGTGGGAGGGTCGACCTCTTGCTCCAGCCCAGATTTCTGCTTCAGTAACTCATGCTT	1620
1621	CCTCTCTCCCCCACCACACCCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGA	1680
1681	AGCTCTGCGTTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCTCAGGAGGTGCCTC	1740
1741	ACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGA	1800
1801	AACATTACTGGGAGGTGGACGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGG	1860
1861	ATGACGTAGACAGGGGAAGAACAATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCC	1920
1921	TCAGACTGACAACAGAACATTTGTATTTACATTCAATCCCCATTTTATCAGCCTCCCCC	1980
1981	CCAGCACCCCTCCTACACGAGTAGGGGTCTTCCTGGACTATGAGGGTGGGACCATCTCCT	2040
2041	TCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGCTGACATGTCAGTTTGAAGGCT	2100
2101	TGTTGAGACCCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCCATATTCA	2160
2161	TATGTCCAGTGTCTTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGCCCCACACCACA	2220

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5C

2221 GACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCTCCGGAG 2280  
2281 CCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCCT 2340  
2341 GAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTG 2400  
2401 GGAGTCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGT 2460  
2461 TTGTGAAAACCTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCA 2520  
2521 TTTGCTAGTCACGGACAGTGATTCTCGCCTCACAGGTGAAGATTAAAGAGACAACGAATG 2580  
2581 TGAATCATGCTTGCAGGTTTGAGGGCCACAGTGTTTGCTAATGGATGTGTTTTATGATT 2640  
2641 ATACATTTTCCCCACCATAAACTCTGTTTGCCTTAATTCCCACATTAATTAACTTTTC 2700  
2701 CTCCTATACCCAAATCCACCCATGGAATAGTTAATTGGAACACCTGCCTTTGTGAGGCTC 2760  
2761 CAAAGAATAAAGAGGAGGTAGGATTTTTCAC TGATTCTATAAGCCCAGCATTACCTGATA 2820  
2821 CCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAAC TACAGGTCCATATCCCTCA 2880  
2881 TTAACACAGACACAAAAATTCTAAATAAAATTTTAACAAAT TAAACTAAACAATATATTT 2940  
2941 AAAGATGATATATAACTACTCAGTGTGGTTTGTCCCACAAAT GCAGAGTTGGTTTAATAT 3000  
3001 TTAAATATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAAAAAAAAAAAAAAAA 3059

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6

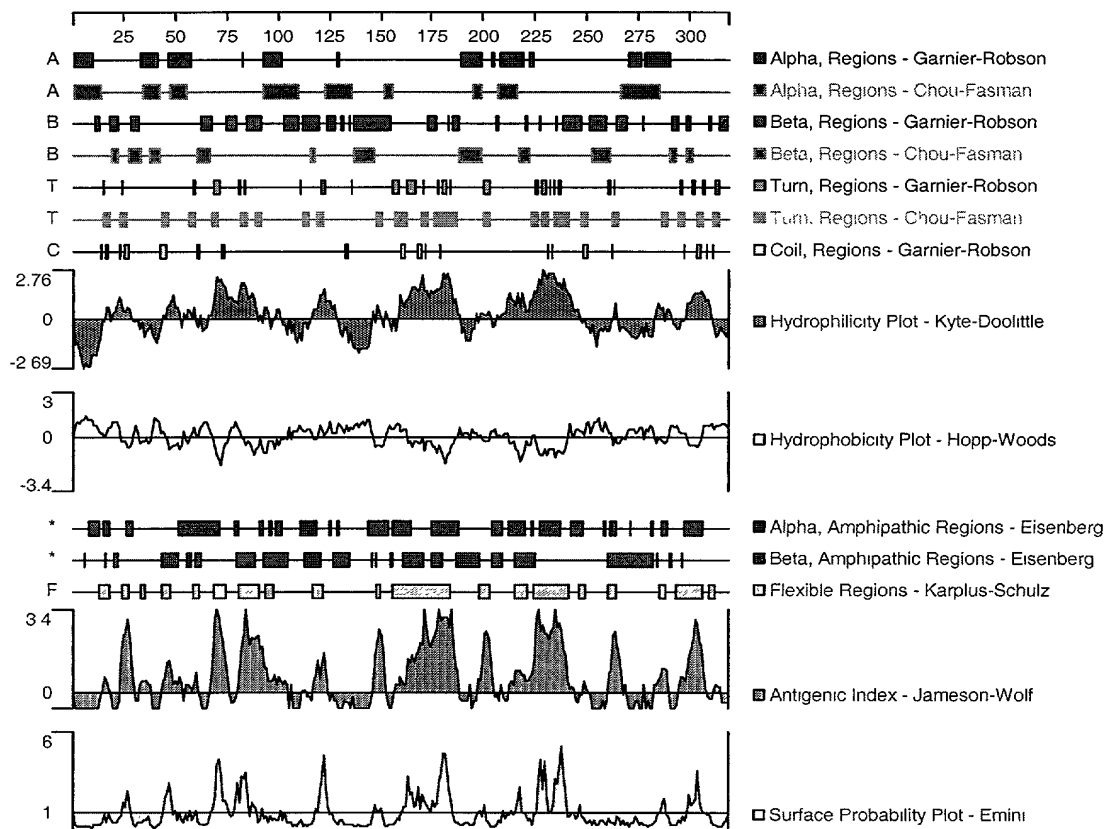


Figure 7A

1 NNCACGAGCCTGTGCCCTGGAAAGGTTGGGAGACTTGGGGGACGACTGGAGAATTGCCAT 60

61 TTGAGGACCAAAGGAGAAAAGAACTACACGCTAATTCTAGAAGGCCTCCTGTCCCTGCC 120

121 TGCTCTGGGTGCTCATGGAACCAGCTGCTGCCCTGCACTTCTCCCGGCCAGCCTCCCTCC 180  
1 M E P A A A L H F S R P A S L L 16

181 TCCTCCTCCTCAGCCTGTGTGCACTGGTCTCAGCCCAGTTTACTGTCTGGGGCCAGCTA 240  
17 L L L S L C A L V S A Q F T V V G P A N 36

241 ATCCCATCCTGGCCATGGTGGGAGAAAACACTACATTACGCTGCCATCTGTCACCCGAGA 300  
37 P I L A M V G E N T T L R C H L S P E K 56

301 AAAATGCTGAGGACATGGAGGTGCGGTGGTTCGGTCTCAGTTCTCCCCCGCAGTGTTCG 360  
57 N A E D M E V R W F R S Q F S P A V F V 76

361 TGTATAAGGGTGGGAGAGAGAGAACAGAGGAGCAGATGGAGGAGTACCGGGAAGAATCA 420  
77 Y K G G R E R T E E Q M E E Y R G R I T 96

421 CCTTTGTGAGCAAAGACATCAACAGGGGCAGCGTGGCCCTGGTCATACATAACGTCACAG 480  
97 F V S K D I N R G S V A L V I H N V T A 116

481 CCCAGGAGAATGGGATCTACCGCTGTTACTTCCAAGAAGGCAGGTCTCTACGATGAGGCCA 540  
117 Q E N G I Y R C Y F Q E G R S Y D E A I 136

541 TCCTACGCCTCGTGGTGGCAGGCCTTGGGTCTAAGCCCCTCATTGAAATCAAGGCCCAAG 600  
137 L R L V V A G L G S K P L I E I K A Q E 156

601 AGGATGGGAGCATCTGGCTGGAGTGCATATCTGGAGGGTGGTACCCAGAGCCCCTCACAG 660  
157 D G S I W L E C I S G G W Y P E P L T V 176

661 TGTGGAGGGACCCCTACGGTGAGGTGTGCCCCCCTGAAGGAGGTTTCCATCGCTGATG 720  
177 W R D P Y G E V V P A L K E V S I A D A 196

721 CTGACGGCCTCTTCATGGTCACCACAGCTGTGATCATCAGAGACAAGTATGTGAGGAATG 780  
197 D G L F M V T T A V I I R D K Y V R N V 216

781 TGTCTGCTCTGTCAACAACACCCTGCTCGGCCAGGAGAAGGAAACTGTCATTTTTATTC 840  
217 S C S V N N T L L G Q E K E T V I F I P 236

841 CAGAATCCTTTATGCCCAGCGCATCTCCCTGGATGGTGGCCCTAGCTGTCATCCTGACCG 900  
237 E S F M P S A S P W M V A L A V I L T A 256

901 CATCTCCCTGGATGGTGTCCATGACTGTCTCCTGGCTGTTTCATCATCTTCATGGCTG 960  
257 S P W M V S M T V I L A V F I I F M A V 276

Figure 7B

961	TCAGCATCTGTTGCATCAAGAACTTCAAAGGGAAAAAAGATTCTGTCAGGGGAAAAAGA	1020
277	S I C C I K K L Q R E K K I L S G E K K	296
1021	AAGTTGAACAAGAGGAAAAAGAAATTGCACAGCAACTTCAAGAAGAATTGCGATGGAGAA	1080
297	V E Q E E K E I A Q Q L Q E E L R W R R	316
1081	GAACATTCTTACATGCTGCTGATGTGGTCTGGATCCAGACACCGCTCATCCCGAGCTCT	1140
317	T F L H A A D V V L D P D T A H P E L F	336
1141	TCCTGTCAGAGGACCGGAGAAGTGTGAGGCGGGCCCCCTACAGGCAGAGAGTGCCTGACA	1200
337	L S E D R R S V R R G P Y R Q R V P D N	356
1201	ACCCAGAGAGATTTCGACAGTCAGCCTTGTGTCTGGGATGGGAGAGCTTCGCCTCAGGGA	1260
357	P E R F D S Q P C V L G W E S F A S G K	376
1261	AACATTACAGGGGAACTTCACAGAGTGGGGACCCACCAGAGCCTATAGAATCAATTCCT	1320
377	H Y R G N F T E W G P T R A Y R I N S L	396
1321	TGGACTCACAGCCATGCAGAAAGCCCTGGCCATCTCAGCAGCCACCGCACAAACCCCTTA	1380
397	D S Q P C R K P W P S Q Q P P H N P P N	416
1381	ATGAAAGACACGCCCTCCTCCCCTCTGGTCACGTAAGAGAACATCTTCCAGCTGCCTTTT	1440
417	E R H A L L P S G H V R E H L P A A F F	436
1441	TCACACCCACTCCAGCCCTCTGCCCCAGTTTTCTCCTCCTCACTAGTCTGTGGCTTTAGT	1500
437	T P T P A L C P S F L L L T S L W L *	455
1501	AGTTCCTTTGCTTGTAATTATGGGATGGGATCCAGGCATAGGGAACCTAGTTGTTTCATAG	1560
1561	CTCCCAGTCAAAAAGAAAGTGAGAGAAGCTGTTGGGCAGCGAACCTACTGTTTAAATCA	1620
1621	GGATAACCAATTAAGCCCAATATGCCAGTTGGCACCAGATGCTGTGGACTTGGAATGAG	1680
1681	GCCAACAGGGTTCACCAGGATGAGAGAGGAGAGGAATCCACAGGACCACCAGAAGGGA	1740
1741	GAGGGAACCAGATATGCAGATCAGAGATAGAGGAAGTGTGAGAGGAAAGGGAGGTCCT	1800
1801	GCTGATTCTCAGAATGGCTTCTGGACCCTGGAGATGTTTGGAACCAATACCGGGCCCT	1860
1861	GTCCTCCCCTGAGAGGATTCTCCCTTTGAAGGAGTCCCTTTGCCGGGTGGGCGTCTTCCT	1920
1921	GGACTATGAAGCTGGAGATGTCTCCTTCTACAACATGAGGGACAGATCACACATCTACAC	1980
1981	ATGTCCCGTTTCAGCCTTTAATGTGCCTGTGAGGCCATTCTTCAGGTTAGGGTCTGATGA	2040

Figure 7C

2041 CAGCCCCATCTTCATCTGCCCTGCACTCACAGGAGCCAGTGGGGTCATGGTGCCTGAAGA 2100  
2101 GGGCCTGAAACTTCACAGAGTGGGGACCCACCAAGGTTGTAAGGATGGCTAAGTCCCACC 2160  
2161 ATAAGAGCTAAAGGGTCCTGGGAGATGATGGCTCATTTCCACCCAACCCAGGATTTCCA 2220  
2221 CAGCACACACCCACAGGCCTGGACCTGGGATGAAGATGAATGAAGAACATGGACTCATGT 2280  
2281 GGATGTGGTTTGGCTCAGATGTCCCTGCAATAAACAAGGGGTCAGTACTTAGTCCCTGAG 2340  
2341 TGTGGTTGAGGTTTGAGGTCCTGGTCGAGCAGGGCAGTACTGGACCAGGTCTACGTCAGC 2400  
2401 ATTCAGGTTCAATGGGGACACCAGTGGCTTCAAACCTCCTGATCTAATTATGTTTTTAGA 2460  
2461 CACTTAGAAGTTATTGAGGACTTTAAAGAACTTTTGTATTATTGGGTAAATATTTATGAC 2520  
2521 ATTTGACCATTGAAACAAAAATTTAAATGTTATCTTTTAATTTATGTTAAAAATAGCATT 2580  
2581 AATAAATCAGTTATAGGTTAATGTAGATAGGATGTTTTGTGAAAAGCAATCTATTGTGT 2640  
2641 CCAAATAAAAAACAAAAAGTGTAATAAAAAAAAAAAAAAAAAAAAA 2682

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 8

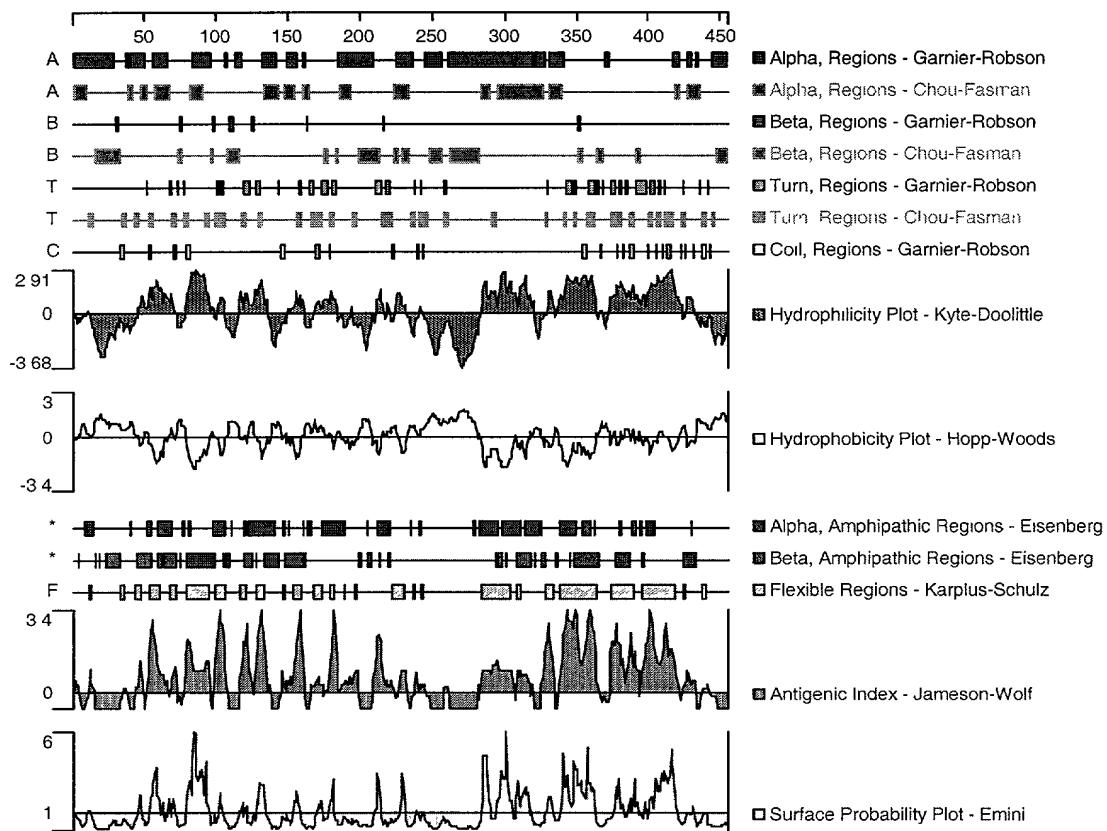




Figure 9A

1	CGATTTCGGCTCCAAACTCCGGCGCTGCAGCCGATCGGACTCTGGGCCGCGGTGGGCACCG	60
61	CGCGCAGCTAGGGAGCCGAGAACC CGCGCGAGCCCCGAGGACGCCAGAGCGCGAGGGTC	120
121	GCTGCGCCTCGCAGAGCCGAGCCGAGTCGAGCCGGCGCCCGGGCTGCCTGGAGACGCC	180
181	GTGACTTTGAAGTGTAACCTCAAGACAGATGGGCGCATGCGGGAGATCGTGTGGTACCGG	240
1	M R E I V W Y R	8
241	GTGACGGATGGTGGCACCATCAAGCAAAAGATCTTCACCTTCGACGCCATGTTCTCCACC	300
9	V T D G G T I K Q K I F T F D A M F S T	28
301	AACTACTCACACATGGAGAACTACCGCAAGCGAGAGGACCTGGTGTACCAGTCCACTGTG	360
29	N Y S H M E N Y R K R E D L V Y Q S T V	48
361	AGGCTGCCCGAGGTCCGGATCTCAGACAATGGTCCCTATGAGTGCCATGTGGGCATCTAC	420
49	R L P E V R I S D N G P Y E C H V G I Y	68
421	GACCGCGCCACCAGGGAGAAGGTGGTCTGGCATCAGGCAACATCTTCCTCAACGTCATG	480
69	D R A T R E K V V L A S G N I F L N V M	88
481	GCTCCTCCACCTCCATTGAAGTGGTGGCTGCTGACACACCAGCCCCCTTCAGCCGCTAC	540
89	A P P T S I E V V A A D T P A P F S R Y	108
541	CAAGCCCAGAACTTCACGCTGGTCTGCATCGTGTCTGGAGGAAAACCAGCACCCATGGTT	600
109	Q A Q N F T L V C I V S G G K P A P M V	128
601	TATTTCAAACGAGATGGGGAACCAATCGACGCAGTGCCCCATCAGAGCCACCAGCTGCC	660
129	Y F K R D G E P I D A V P L S E P P A A	148
661	AGCTCCGGCCCCCTACAGGACAGCAGGCCCTTCCGCAGCCTTCTGCACCGTGACCTGGAT	720
149	S S G P L Q D S R P F R S L L H R D L D	168
721	GACACCAAGATGCAGAAGTCACTGTCCCTCCTGGACGCCGAGAACC GGGGTGGGCGACCC	780
169	D T K M Q K S L S L L D A E N R G G R P	188
781	TACACGAGCGCCCCCTCCCGTGGCCTGACCCAGATCCCAACATCCTCCTCCAGCCAACC	840
189	Y T E R P S R G L T P D P N I L L Q P T	208
841	ACAGAGAACATACCAGAGACGGTCGTGAGCCGTGAGTTTCCCCGCTGGGTCCACAGCGCC	900
209	T E N I P E T V V S R E F P R W V H S A	228
901	GAGCCCACCTACTTCTGCGCCACAGCCGCACCCCGAGCAGTGACGGCACTGTGGAAGTA	960
229	E P T Y F L R H S R T P S S D G T V E V	248
961	CGTGCCCTGCTCACCTGGACCCCTCAACCCACAGATCGACAACGAGGCCCTCTTCAGCTGC	1020
249	R A L L T W T L N P Q I D N E A L F S C	268

Figure 9B

1021	GAGGTCAAGCACCCAGCTCTGTTCGATGCCCATGCAGGCAGAGGTCACGCTGGTTGCCCC	1080
269	E V K H P A L S M P M Q A E V T L V A P	288
1081	AAAGGACCCAAAATTGTGATGACGCCCAGCAGAGCCCGGTAGGGGACACAGTGAGGATT	1140
289	K G P K I V M T P S R A R V G D T V R I	308
1141	CTGGTCCATGGGTTTCAGAACGAAGTCTTCCCGAGCCCATGTTACGTGGACGCGGGTT	1200
309	L V H G F Q N E V F P E P M F T W T R V	328
1201	GGGAGCCGCCTCCTGGACGGCAGCGCTGAGTTCGACGGGAAGGAGCTGGTGCTGGAGCGG	1260
329	G S R L L D G S A E F D G K E L V L E R	348
1261	GTTCCCGCCGAGCTCAATGGCTCCATGTATCGCTGCACCGCCAGAACCCACTGGGCTCC	1320
349	V P A E L N G S M Y R C T A Q N P L G S	368
1321	ACCGACACGCACACCCGGCTCATCGTGTTTGAAAACCCAAATATCCCAAGAGGAACGGAG	1380
369	T D T H T R L I V F E N P N I P R G T E	388
1381	GACTCTAATGGTTCCATTGGCCCCACTGGTGCCCGGCTCACCTTGGTGCTCGCCCTGACA	1440
389	D S N G S I G P T G A R L T L V L A L T	408
1441	GTGATTCTGGAGCTGACGTGAAGGCACCCGCCCCCGGCCACTCCATCAGGCACTGACATCT	1500
409	V I L E L T *	415
1501	CCGCGACCGGTTTTTCATTTCTTTTCTAAACTATTTCCAGTCTTGTTCTTAGTCTCTTTCC	1560
1561	ATCTGTGTCTTGGCTTCTTCAGTCGGTTTAATTAAAACAAACAGAACAATTTTCCCCACA	1620
1621	AA	1680
1681	AA 1724	

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 10

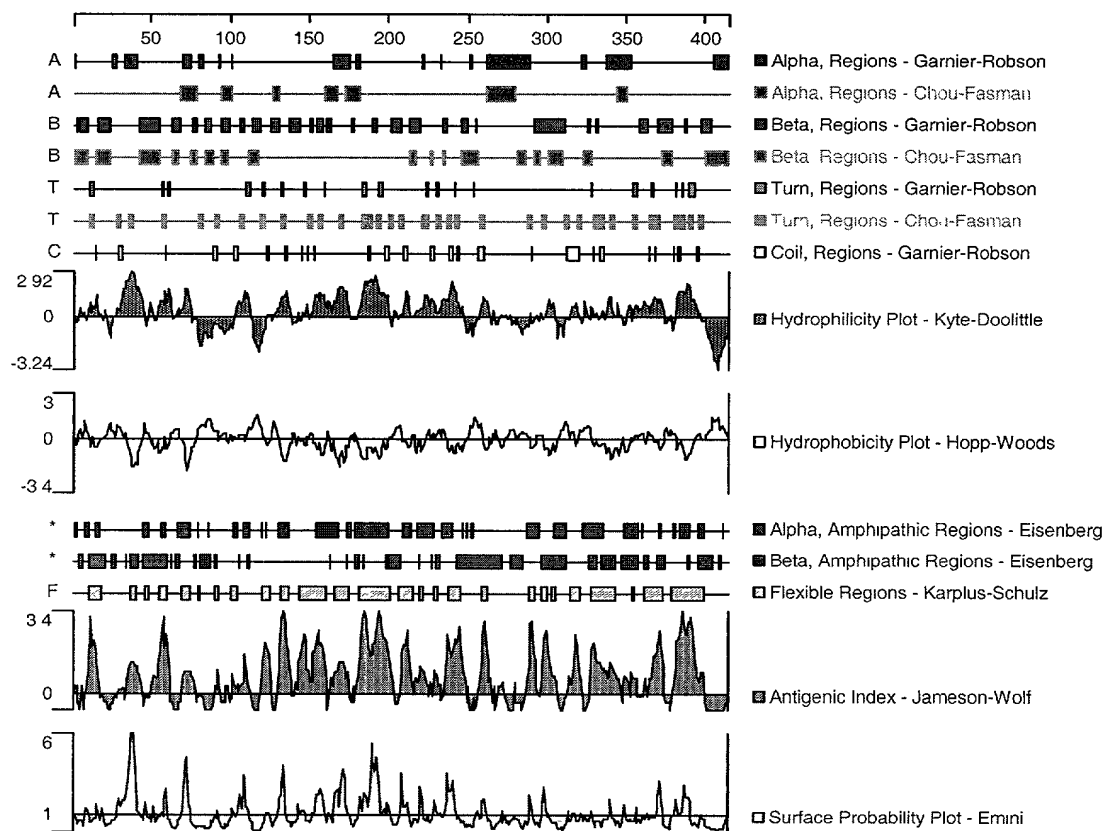


Figure 11

```

1  CACGAGCCTGTGCCCCGTGAAAGGTTGGAGACTTGGGGGACGACTGGAGAATTGCCATTT 60
61  GAGGACCAAAGGAGAAAAGAACTACACGCTAATTCTAGAAGGCCTCCTGTCCCTGCCTG 120
121 CTCTGGGTGCTCATGGAACCAGCTGCTGCCCTGCACTTCTCCCGGCCAGCCTCCCTCCTC 180
1  M E P A A A L H F S R P A S L L 16
181 CTCCTCCTCAGCCTGTGTGCACTGGTCTCAGCCCAGGTCAGTGTCTGCGTGGGGCCCACTGAT 240
17 L L L S L C A L V S A Q V T V V G P T D 36
241 CCCATCCTGGCCATGGTGGGAGAAAACACTACGTTACGATGCTGTCTGTACCCGAGGAA 300
37 P I L A M V G E N T T L R C C L S P E E 56
301 AATGCTGAGGACATGGAGGTGCGGTGGTTCCAGTCTCAGTTCTCCCCTGCAGTGTTTGTG 360
57 N A E D M E V R W F Q S Q F S P A V F V 76
361 TATAAGGGTGGAAAGAGAGAGAACAGAGGAGCAGAAGGAGGAGTACCGAGGGAGAACCACC 420
77 Y K G G R E R T E E Q K E E Y R G R T T 96
421 TTTGTGAGCAAAGACAGCAGGGGCGAGCGTGGCCCTGATCATAACAATGTCACAGCCGAG 480
97 F V S K D S R G S V A L I I H N V T A E 116
481 GATAACGGCATCTACCAAGTGTACTTCCAAGAAGGCAGGTCCTGCAATGAGGCCATCCTG 540
117 D N G I Y Q C Y F Q E G R S C N E A I L 136
541 CACCTTGTGGTGGCAGACCAGCACAAATCCTCTTTCTGATCCCCATTCGCAGGGGACA 600
137 H L V V A D Q H N P L S W I P I P Q G T 156
601 CTCTCCCTATGAAAAGAAGATTCCAGGGGAAAAATCCTTCTCCTGCACAAGGGCCACCA 660
157 L S L * 160
661 TGAGTGAGTTTGCCTGCTAAGCCGTGGGCTTGACTTCTTGAGAAGCACATGCAGAACTC 720
721 AGTTGAGGCCATGAGCCGGGGGAAAATGGTGAATCTCGGAAGAGAAGTCCTATGCCTGCC 780
781 TTAGCACTGAGCTGTGCACTTCTGAGAGTGAGAGGAGACACCATCAATAATTGTCTTGGG 840
841 ACAACTGGAATAAACAGTGACTGCCCAGAGAACTACGATATTTGAAATCTTATTTCTTGA 900
901 TGAATATTCATCCTGACTTCTTTCTGAAATGCTGTTTGCAAAGAGAGTGACTTATATGT 960
961 AAGTAGAGCGTTTATTATTAAGCAAGACTTAATACAGAAGCAAAAAAAAAAAAAAAAAAAAA 1019

```

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 12

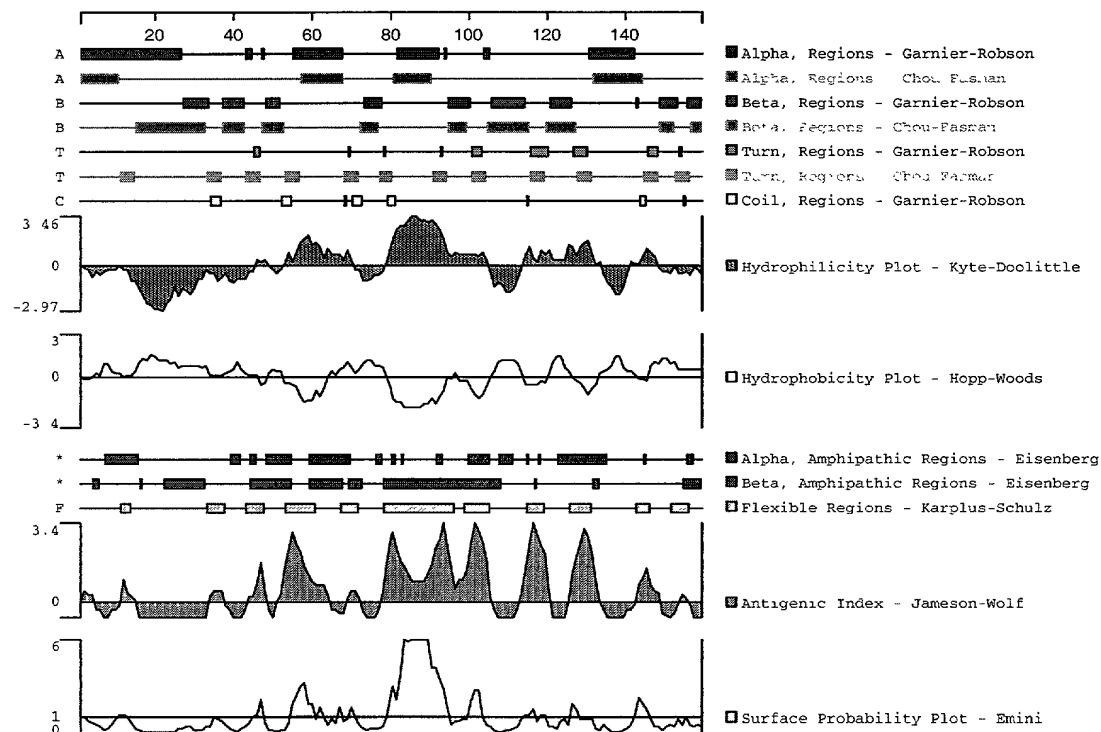


Figure 13A

1	ACATCCATGGCTCTAATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAG	60
1	M A L M L S L V L S L L K L G S G Q	18
61	TGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGAGGACGCAGCATTC	120
19	W Q V F G P D K P V Q A L V G E D A A F	38
121	TCCTGTTTCCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGC	180
39	S C F L S P K T N A E A M E V R F F R G	58
181	CAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATG	240
59	Q F S S V V H L Y R D G K D Q P F M Q M	78
241	CCACAGTATCAAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCT	300
79	P Q Y Q G R T K L V K D S I A E G R I S	98
301	CTGAGGCTGGAAAACATTACTGTGTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCC	360
99	L R L E N I T V L D A G L Y G C R I S S	118
361	CAGTCTTACTACCAGAAGGCCATCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCCT	420
119	Q S Y Y Q K A I W E L Q V S A L G S V P	138
421	CTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTACTCTGTTCAGTCCCTCGGGC	480
139	L I S I T G Y V D R D I Q L L C Q S S G	158
481	TGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTCCACAGAC	540
159	W F P R P T A K W K G P Q G Q D L S T D	178
541	TCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAA	600
179	S R T N R D M H G L F D V E I S L T V Q	198
601	GAGAACGCCGGGAGCATATCCTGTTCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAA	660
199	E N A G S I S C S M R H A H L S R E V E	218
661	TCCAGGGTACAGATAGGAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAA	720
219	S R V Q I G D T F F E P I S W H L A T K	238
721	GTACTGGGAATACTCTGCTGTGGCCTATTTTTTGGCATTGTTGGACTGAAGATTTTCTTC	780
239	V L G I L C C G L F F G I V G L K I F F	258
781	TCCAAATTCAGTGGAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCA	840
259	S K F Q W K I Q A E L D W R R K H G Q A	278
841	GAATTGAGAGACGCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCAC	900
279	E L R D A R K H A V E V T L D P E T A H	298
901	CCGAAGCTCTGCGTTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCCAGGAGGTG	960
299	P K L C V S D L K T V T H R K A P Q E V	318



Figure 14

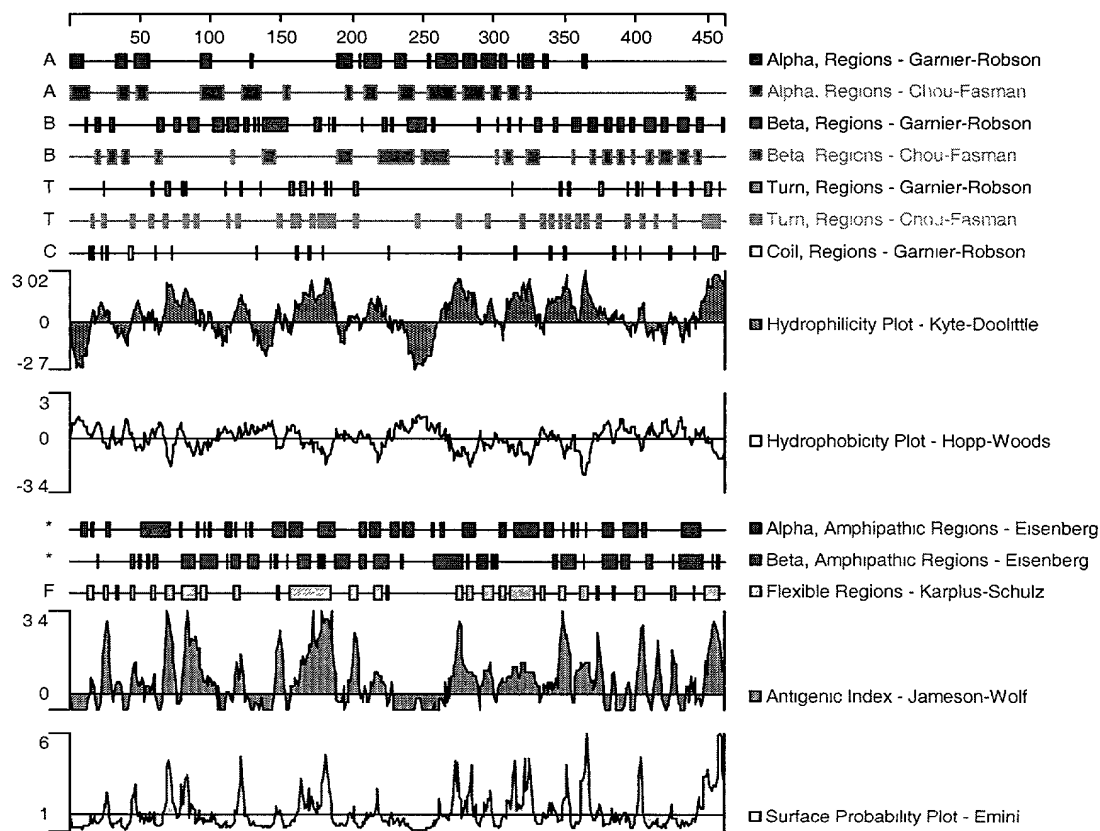




Figure 15

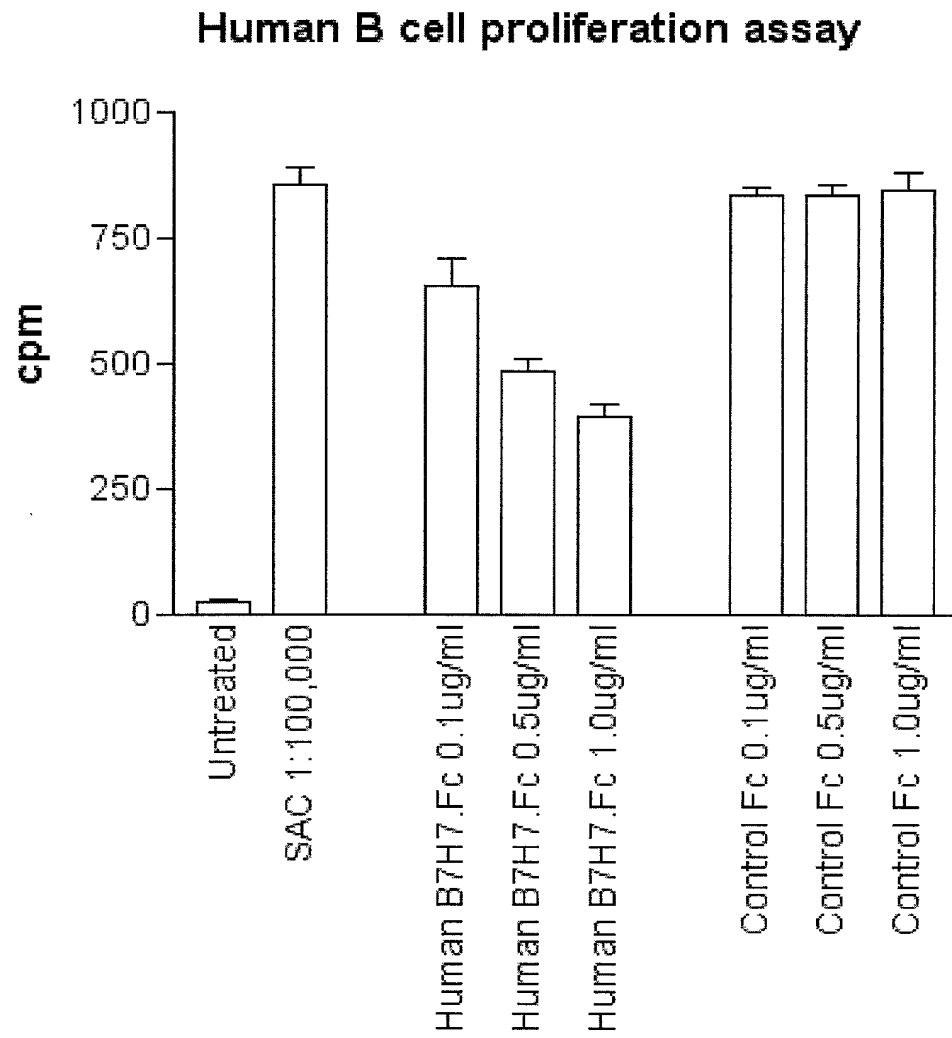


Figure 16

